Work relations for time-dependent states

Naoko Nakagawa[1] and Shin-ichi Sasa[2]

[1] *College of Science, Ibaraki University, Mito, Ibaraki 310-8512, Japan*
[2] *Department of Physics, Kyoto University, Kyoto 606-8502, Japan*
(Dated: November 28, 2012)

For time-dependent states generated by an external operation, a generalized free energy may be introduced by the relative entropy with respect to an equilibrium state realized after sufficient relaxation from the time-dependent states. Recently, by studying over-damped systems, Sivak and Crooks presented a new formula that relates the generalized free energy with measurable thermodynamic works. We re-derive this relation with emphasizing a connection to an extended Clausius relation proposed in the framework of steady state thermodynamics. As a natural consequence, we generalize this relation to be valid for systems with momentum degrees of freedom, where the Shannon entropy in the generalized free energy is replaced by a symmetric one.

Introduction: Over the last two decades scientific activities in non-equilibrium statistical mechanics have discovered universal relations regarded as extensions of the second law of thermodynamics. Examples of such universal relations include fluctuation theorems that claim identities leading to the positivity of entropy production in non-equilibrium steady states (NESS) [1–6], Jarzynski's work relations by which free energy differences are estimated in non-equilibrium processes [7–10], and steady state thermodynamics (SST) in which state variables are defined by (excess) heat or works in NESS [11–23].

Recently, for over-damped systems near equilibrium, Sivak and Crooks presented a work relation associated with a generalized free energy for a time-dependent state realized by a finite-speed operation to an equilibrium state [24]. The free energy they defined is expressed as the relative entropy of the probability density with respect to an equilibrium state realized after sufficient relaxation [25]. Remarkably, when the quasi-static limit is considered, a variation of the Sivak and Crooks relation (see (29)) becomes equivalent to a thermodynamic relation between work and the difference in equilibrium free energies. This means that the Sivak and Crooks relation is an extension of the second law of thermodynamics.

A natural question arising here is whether the Sivak and Crooks relation is connected to previously known relations, such as fluctuation theorems, Jarzynski's work relations, or formulas proposed in SST. The purpose of this Rapid Communication is to find a direct connection to previous studies by presenting a simpler derivation of the Sivak and Crooks relation. Our re-derivation clearly demonstrates that the Sivak and Crooks relation corresponds to a non-stationary version of the extended Clausius relation proposed in SST [17, 18]. Following the idea proposed in SST, we successfully generalize the Sivak and Crooks relation to be valid for systems with momentum degrees of freedom.

Preliminaries: We study systems consisting of N particles under isothermal environment of inverse temperature β. Examples of such systems include (i) a Hamiltonian system in contact with a heat bath, (ii) an under-damped Langevin equation, or (iii) an over-damped Langevin equation. The microscopic coordinates of N particles are collectively denoted by $\Gamma = (\mathbf{r}_1, \ldots, \mathbf{r}_N; \mathbf{p}_1, \ldots, \mathbf{p}_N)$ for Hamiltonian systems and under-damped Langevin equations or $\Gamma = (\mathbf{r}_1, \ldots, \mathbf{r}_N)$ for over-damped Langevin equations. For all the considered cases, a Hamiltonian $H_\nu(\Gamma)$ characterizes the system, where ν represents a set of parameters. For simplicity, we assume

$$H_\nu(\Gamma) = H_\nu(\Gamma^*), \tag{1}$$

where Γ^* represents the time-reversal of Γ defined as $\Gamma^* = (\mathbf{r}_1, \ldots, \mathbf{r}_N; -\mathbf{p}_1, \ldots, -\mathbf{p}_N)$. (Note that $\Gamma^* = \Gamma$ for over-damped cases.) We also assume that an external agent operates the system according to a prefixed protocol $\hat{\nu}$ specified by a function $\nu(t)$ of $t \in [0, \tau]$. We fix a function $\bar{\nu}(t)$ of $t \in [0, 1]$ with $\bar{\nu}(0) = \nu_0$ and $\bar{\nu}(1) = \nu_1$, by which we have $\nu(t) = \bar{\nu}(\epsilon t)$ with $\tau = \epsilon^{-1}$, where the typical speed of the protocol is denoted as ϵ. For any protocol $\hat{\nu}$, we define its time reversed protocol $\hat{\nu}^\dagger$ by $\nu^\dagger(t) = \bar{\nu}(1 - \epsilon t)$.

Let $\hat{\Gamma} = (\Gamma(t))_{t \in [0, \tau]}$ be a path in a time interval $[0, \tau]$, and $\hat{\Gamma}^\dagger = (\Gamma^*(\tau - t))_{t \in [0, \tau]}$ be the time reversed path of $\hat{\Gamma}$. For the system with $\hat{\nu}$, we denote the probability density of $\hat{\Gamma}$ as $\mathcal{T}_\epsilon(\hat{\Gamma})$, provided that $\Gamma(0)$ is given. Note that the $\bar{\nu}$ dependence of $\mathcal{T}$ is not written explicitly, but the ϵ dependence is shown. For the reversed protocol $\hat{\nu}^\dagger$, we express the probability density of $\hat{\Gamma}$ as $\mathcal{T}_\epsilon^\dagger(\hat{\Gamma})$. The important assumption of our model is that $\mathcal{T}_\epsilon(\hat{\Gamma})$ satisfies the local detailed balance condition

$$\mathcal{T}_\epsilon(\hat{\Gamma}) \, e^{\beta Q(\hat{\Gamma})} = \mathcal{T}_\epsilon^\dagger(\hat{\Gamma}^\dagger), \tag{2}$$

where $Q(\hat{\Gamma})$ represents the energy that flows from the heat bath to the system in the path $\hat{\Gamma} = (\Gamma(t))_{t \in [0, \tau]}$ for the protocol $\hat{\nu}$. In fact, (2) is valid for both over-damped and under-damped Langevin equations with the

energetic interpretation proposed in Refs. [26, 27]. For a mechanical description of a heat bath in contact with a Hamiltonian system, (2) is derived in Ref. [28]. Throughout this Rapid Communication, the Boltzmann constant is set to be unity.

When the parameter value is fixed as ν (that is, $\bar{\nu}(t)$ is independent of t), the probability density converges to the canonical one

$$\rho_\nu^{\text{can}}(\Gamma) = e^{\beta(F_\nu - H_\nu(\Gamma))} \tag{3}$$

in the limit $t \to \infty$, where F_ν is the free energy defined by

$$F_\nu = -\beta^{-1} \log \left(\int d\Gamma e^{-\beta H_\nu(\Gamma)} \right). \tag{4}$$

Now, we fix a functional form of the scaled protocol $\bar{\nu}(t)$ ($0 \le t \le 1$) and assume that the system is in equilibrium at $t = 0$ with $\nu(0) = \nu_0$. Let $\rho_\epsilon(\Gamma)$ be the probability density of the time-dependent state at the end of the protocol $\hat{\nu}$. Formally, in terms of the path probability measure $\mathcal{D}\hat{\Gamma}$, we express $\rho_\epsilon(\Gamma)$ as

$$\rho_\epsilon(\Gamma) = \int \mathcal{D}\hat{\Gamma} \, \rho_{\nu_0}^{\text{can}}(\Gamma(0))\mathcal{T}_\epsilon(\hat{\Gamma})\delta(\Gamma(\tau) - \Gamma). \tag{5}$$

We shall characterize this probability density.

Let us suppose that the value of ν in the system is fixed as $\nu(\tau) = \nu_1$. The system relaxes from the time-dependent state, and the probability density approaches the canonical one $\rho_{\nu_1}^{\text{can}}$. We then measure the distance of ρ_ϵ from $\rho_{\nu_1}^{\text{can}}$ by the relative entropy $\mathcal{D}(\rho_\epsilon|\rho_{\nu_1}^{\text{can}})$, whose definition is given as

$$\mathcal{D}(\varphi|\varphi') := \int d\Gamma \varphi(\Gamma) \log \frac{\varphi(\Gamma)}{\varphi'(\Gamma)} \tag{6}$$

for any two probability densities $\varphi(\Gamma)$ and $\varphi'(\Gamma)$. The relative entropy has been a useful quantity in the characterization of the dissipation [29] as well as in information science [30]. In particular, it should be noted that $\mathcal{D}(\rho_\epsilon|\rho_{\nu_1}^{\text{can}})$ is connected to a generalized free energy $\mathcal{F}(\rho_\epsilon)$ defined as

$$\mathcal{F}(\rho_\epsilon) = \mathcal{U}(\rho_\epsilon) - \beta^{-1}\mathcal{S}(\rho_\epsilon), \tag{7}$$

with

$$\mathcal{U}(\rho_\epsilon) = \int d\Gamma \rho_\epsilon(\Gamma) H_\nu(\Gamma), \tag{8}$$

$$\mathcal{S}(\rho_\epsilon) = -\int d\Gamma \rho_\epsilon(\Gamma) \log \rho_\epsilon(\Gamma). \tag{9}$$

The calligraphic font used in $\mathcal{D}$, $\mathcal{F}$, $\mathcal{U}$, and $\mathcal{S}$ represents their functional dependence of ρ_ϵ. That is, $\mathcal{S}$, $\mathcal{F}$, and $\mathcal{U}$ are not state functions uniquely determined for (β, ν), and $\mathcal{F}(\rho_\nu^{\text{can}}) = F_\nu$ is a function of (β, ν). By direct calculation, it is confirmed that [25]

$$\beta(\mathcal{F}(\rho_\epsilon) - F_{\nu_1}) = \mathcal{D}(\rho_\epsilon|\rho_{\nu_1}^{\text{can}}). \tag{10}$$
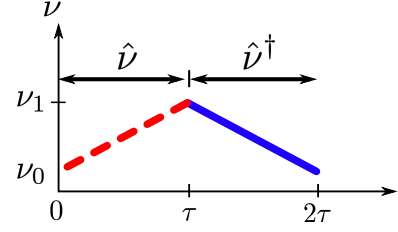

FIG. 1: (Color online) Schematic of the protocol in (17). The reverse protocol $\hat{\nu}^\dagger$ (solid blue line) is performed immediately after the protocol $\hat{\nu}$ (dashed red line). The work W in $\hat{\nu}$ and $W^\dagger$ in $\hat{\nu}^\dagger$ are measured.

For later convenience, we introduce the following notations. First, for any function $f(\hat{\Gamma})$ of $\hat{\Gamma}$, we define conditional expectations

$$\langle f \rangle_{\Gamma \to} = \int \mathcal{D}\hat{\Gamma} \, \delta(\Gamma(0) - \Gamma)\mathcal{T}_\epsilon(\hat{\Gamma}) \, f(\hat{\Gamma}) \tag{11}$$

with a fixed initial state Γ, and

$$\langle f \rangle_{\rho_{\nu_0}^{\text{can}} \to \Gamma} = \int \mathcal{D}\hat{\Gamma} \frac{\mathcal{T}_{\hat{\nu}}[\hat{\Gamma}]\delta(\Gamma(\tau) - \Gamma) \, f(\hat{\Gamma})\rho_{\nu_0}^{\text{can}}(\Gamma(0))}{\rho_\epsilon(\Gamma)} \tag{12}$$

with a fixed final state Γ, starting from the initial equilibrium state $\rho_{\nu_0}^{\text{can}}$. Similarly, we define

$$\langle f \rangle_{\rho \to} := \int \mathcal{D}\hat{\Gamma}\rho(\Gamma(0))\mathcal{T}_\epsilon[\hat{\Gamma}]f(\hat{\Gamma}), \tag{13}$$

$$\langle f \rangle_{\rho \to}^\dagger := \int \mathcal{D}\hat{\Gamma}\rho(\Gamma(0))\mathcal{T}_\epsilon^\dagger[\hat{\Gamma}]f(\hat{\Gamma}), \tag{14}$$

which are measured in the protocol $\hat{\nu}$ and its reversed protocol $\hat{\nu}^\dagger$, respectively, for an initial distribution ρ.

Sivak and Crooks relation: Let $W(\hat{\Gamma})$ and $W^\dagger(\hat{\Gamma})$ be the work done by an external agent with protocols $\hat{\nu}$ and $\hat{\nu}^\dagger$ in the path $\hat{\Gamma}$, respectively. That is,

$$W(\hat{\Gamma}) := \int_0^\tau dt \left. \frac{\partial H_{\nu(t)}(\Gamma(s))}{\partial t} \right|_{s=t}, \tag{15}$$

$$W^\dagger(\hat{\Gamma}) := \int_0^\tau dt \left. \frac{\partial H_{\nu^\dagger(t)}(\Gamma(s))}{\partial t} \right|_{s=t}. \tag{16}$$

Note that $W^\dagger(\hat{\Gamma}^\dagger) = -W(\hat{\Gamma})$ along the time reversed path $\hat{\Gamma}^\dagger$. For over-damped cases, Sivak and Crooks have derived the following relation [24]:

$$\mathcal{F}(\rho_\epsilon) - F_{\nu_1} = \frac{1}{2}\left(\langle W^\dagger \rangle_{\rho_{\nu_1}^{\text{can}} \to}^\dagger - \langle W^\dagger \rangle_{\rho_\epsilon \to}^\dagger \right) + O(\epsilon^3) \tag{17}$$

In order to obtain $\langle W^\dagger \rangle_{\rho_\epsilon \to}^\dagger$ in experiments, we first perform the protocol $\hat{\nu}$, and then start the protocol $\hat{\nu}^\dagger$ exactly when $\hat{\nu}$ finishes (without pause). See Fig. 1. Below, we re-derive (17) by using a simpler procedure.

Fundamental relations: By using the energy conservation

$$H_{\nu_1}(\Gamma(\tau)) - H_{\nu_0}(\Gamma(0)) = Q(\hat{\Gamma}) + W(\hat{\Gamma}), \qquad (18)$$

we rewrite the local detailed balance condition (2) as

$$\rho_{\nu_0}^{\mathrm{can}}(\Gamma(0))\mathcal{T}_\epsilon[\hat{\Gamma}]e^{-\beta\frac{W_{\mathrm{irr}}(\hat{\Gamma})}{2}} = \rho_{\nu_1}^{\mathrm{can}}(\Gamma^*(\tau))\mathcal{T}_\epsilon^\dagger[\hat{\Gamma}^\dagger]e^{-\beta\frac{W_{\mathrm{irr}}^\dagger(\hat{\Gamma}^\dagger)}{2}} \quad (19)$$

where $W_{\mathrm{irr}}(\hat{\Gamma}) = W(\hat{\Gamma}) - \Delta F$ and $W_{\mathrm{irr}}^\dagger(\hat{\Gamma}^\dagger) = W^\dagger(\hat{\Gamma}^\dagger) + \Delta F = -W_{\mathrm{irr}}(\hat{\Gamma})$ with $\Delta F = F_{\nu_1} - F_{\nu_0}$. Integrating (19) over all paths, we have a symmetric Jarzynski-type equality

$$\left\langle e^{-\beta\frac{W_{\mathrm{irr}}}{2}}\right\rangle_{\rho_{\nu_0}^{\mathrm{can}}\to} = \left\langle e^{-\beta\frac{W_{\mathrm{irr}}^\dagger}{2}}\right\rangle_{\rho_{\nu_1}^{\mathrm{can}}\to}^\dagger. \qquad (20)$$

Furthermore, by integrating (19) over paths $\hat{\Gamma}$ that satisfy the condition $\Gamma(\tau) = \Gamma$, we obtain the following expression for ρ_ϵ:

$$\rho_\epsilon(\Gamma) = \rho_{\nu_1}^{\mathrm{can}}(\Gamma)\frac{\left\langle e^{-\beta\frac{W_{\mathrm{irr}}^\dagger}{2}}\right\rangle_{\Gamma^*\to}^\dagger}{\left\langle e^{-\beta\frac{W_{\mathrm{irr}}}{2}}\right\rangle_{\rho_{\nu_0}^{\mathrm{can}}\to\Gamma}}. \qquad (21)$$

The relations (20) and (21) are basic equations in our derivation.

To expand the relations (20) and (21) in terms of ϵ, we prepare the following three estimations. First, since $W_{\mathrm{irr}}(\hat{\Gamma}) \to 0$ for typical $\hat{\Gamma}$ in the quasi-static limit $\epsilon \to 0$, we assume

$$W_{\mathrm{irr}}(\hat{\Gamma}) = O(\epsilon) \qquad (22)$$

for typical $\hat{\Gamma}$. Second, for any quantity $f(\hat{\Gamma}) \simeq O(\epsilon^0)$, (19) yields

$$\langle f\rangle_{\rho_{\nu_0}^{\mathrm{can}}\to} = \langle f^*\rangle_{\rho_{\nu_1}^{\mathrm{can}}\to}^\dagger + O(\epsilon), \qquad (23)$$

$$\langle f\rangle_{\Gamma\to} = \langle f^*\rangle_{\rho_{\nu_0}^{\mathrm{can}}\to\Gamma^*}^\dagger + O(\epsilon), \qquad (24)$$

where $f^*(\hat{\Gamma}) := f(\hat{\Gamma}^\dagger)$.

Now, by setting $f = \epsilon^{-2}W_{\mathrm{irr}}W_{\mathrm{irr}}$ in (23), we obtain $\langle W_{\mathrm{irr}}; W_{\mathrm{irr}}\rangle_{\rho_{\nu_0}^{\mathrm{can}}\to} = \langle W_{\mathrm{irr}}^\dagger; W_{\mathrm{irr}}^\dagger\rangle_{\rho_{\nu_1}^{\mathrm{can}}\to}^\dagger + O(\epsilon^3)$, where $\langle A; B\rangle = \langle AB\rangle - \langle A\rangle\langle B\rangle$. Thus, the expansion of (20) leads to

$$\langle W_{\mathrm{irr}}\rangle_{\rho_{\nu_0}^{\mathrm{can}}\to} = \langle W_{\mathrm{irr}}^\dagger\rangle_{\rho_{\nu_1}^{\mathrm{can}}\to}^\dagger + O(\epsilon^3). \qquad (25)$$

Similarly, by using (24) in the expansion of (21), we derive

$$\rho_\epsilon(\Gamma) = \rho_{\nu_1}^{\mathrm{can}}(\Gamma)\exp\left[\frac{\beta}{2}\left(\langle W_{\mathrm{irr}}\rangle_{\rho_{\nu_0}^{\mathrm{can}}\to\Gamma} - \langle W_{\mathrm{irr}}^\dagger\rangle_{\Gamma^*\to}^\dagger\right)\right] + O(\epsilon^3). \qquad (26)$$

A similar expression was proposed in NESS [31, 32].

Derivation of Sivak and Crooks relation: In this paragraph, we consider over-damped cases ($\Gamma^* = \Gamma$). By substituting (26) into (10), we obtain

$$\mathcal{F}(\rho_\epsilon) - F_{\nu_1} = \frac{1}{2}\left(\langle W_{\mathrm{irr}}\rangle_{\rho_{\nu_0}^{\mathrm{can}}\to} - \langle W_{\mathrm{irr}}^\dagger\rangle_{\rho_\epsilon\to}^\dagger\right) + O(\epsilon^3). \qquad (27)$$

Further, by substituting (25) into (27), and by using the trivial identity

$$\langle W_{\mathrm{irr}}\rangle_{\rho_{\nu_1}^{\mathrm{can}}\to}^\dagger - \langle W_{\mathrm{irr}}\rangle_{\rho_\epsilon\to}^\dagger = \langle W^\dagger\rangle_{\rho_{\nu_1}^{\mathrm{can}}\to}^\dagger - \langle W^\dagger\rangle_{\rho_\epsilon\to}^\dagger, \qquad (28)$$

we arrive at the Sivak and Crooks relation (17).

Here, by noting $\Delta F = F_{\nu_1} - F_{\nu_0}$, (27) is further rewritten as

$$\mathcal{F}(\rho_\epsilon) - F_{\nu_0} = \frac{1}{2}\left(\langle W\rangle_{\rho_{\nu_0}^{\mathrm{can}}\to} - \langle W^\dagger\rangle_{\rho_\epsilon\to}^\dagger\right) + O(\epsilon^3). \qquad (29)$$

By combining this expression with (7) and (18), we can derive the following relation for the entropy:

$$\mathcal{S}(\rho_\epsilon) - S_{\nu_0} = \frac{\beta}{2}\left(\langle Q\rangle_{\rho_{\nu_0}^{\mathrm{can}}\to} - \langle Q^\dagger\rangle_{\rho_\epsilon\to}^\dagger\right) + O(\epsilon^3). \qquad (30)$$

This is an extended Clausius relation, because (30) leads to the standard equilibrium Clausius relation

$$S_{\nu_1} - S_{\nu_0} = \beta\langle Q\rangle_{\rho_{\nu_0}^{\mathrm{can}}\to} \qquad (31)$$

in the limit $\epsilon \to 0$. Similarly, (29) leads to the standard equilibrium work relation in the limit $\epsilon \to 0$.

Let us compare (30) with the extended Clausius relation in SST, where it has been shown that heat measurement enables us to experimentally determine state variables for non-equilibrium steady states driven by a non-conservative force or by non-equilibrium boundary conditions [17]. It is found that (30) takes the same form as an expression reported in Ref. [17]. (See eq. (11) of Ref. [17].) This is not accidental. In fact, the derivation method of (30) or (29) is essentially the same as that used in Ref. [17]. Furthermore, the Sivak and Crooks relations can be easily formulated for time-dependent states generated by an external operation to NESS. From these considerations, in our viewpoint, (30) corresponds to a non-stationary version of the extended Clausius relation in SST.

Under-damped cases: Next, we consider under-damped systems ($\Gamma^* \neq \Gamma$). Following the idea proposed in Ref. [17], we define the symmetric Shannon entropy

$$\mathcal{S}^{\mathrm{sym}}(\rho_\epsilon) = -\int d\Gamma\,\rho_\epsilon(\Gamma)\log\sqrt{\rho_\epsilon(\Gamma)\rho_\epsilon(\Gamma^*)} \qquad (32)$$

and the corresponding free energy

$$\mathcal{F}^{\mathrm{sym}}(\rho_\epsilon) = \mathcal{U}(\rho_\epsilon) - \beta^{-1}\mathcal{S}^{\mathrm{sym}}(\rho_\epsilon). \qquad (33)$$

It should be noted that $\mathcal{S} = \mathcal{S}^{\mathrm{sym}}$ and $\mathcal{F} = \mathcal{F}^{\mathrm{sym}}$ when $\Gamma = \Gamma^*$. In Ref. [17], the symmetric Shannon entropy

was proposed as a new state variable determined operationally by measuring the excess heat in the quasi-static operation connecting two NESS near equilibrium. Explicitly, the symmetric free energy $\mathcal{F}^{\mathrm{sym}}$ is written as

$$\beta(\mathcal{F}^{\mathrm{sym}}(\rho_\epsilon) - F_{\nu_1}) = \int d\Gamma \rho_\epsilon(\Gamma) \log \frac{\sqrt{\rho_\epsilon(\Gamma)\rho_\epsilon(\Gamma^*)}}{\rho_{\nu_1}^{\mathrm{can}}(\Gamma)} \quad (34)$$

The right-hand side is not the relative entropy. Rather, it can be called "symmetric relative entropy."

Now, by substituting (26) into the right-hand side of (34) and by extracting the term $\langle W_{\mathrm{irr}}\rangle_{\rho_{\nu_0}^{\mathrm{can}}\to\Gamma} - \langle W_{\mathrm{irr}}^\dagger\rangle_{\Gamma\to}^\dagger$, we write

$$\log \frac{\rho_\epsilon(\Gamma)}{\rho_{\nu_1}^{\mathrm{can}}(\Gamma)} + \log \frac{\rho_\epsilon(\Gamma^*)}{\rho_{\nu_1}^{\mathrm{can}}(\Gamma^*)}$$
$$= \beta\left(\langle W_{\mathrm{irr}}\rangle_{\rho_{\nu_0}^{\mathrm{can}}\to\Gamma} - \langle W_{\mathrm{irr}}^\dagger\rangle_{\Gamma\to}^\dagger\right)$$
$$- \frac{\beta}{2}\left(\langle W_{\mathrm{irr}}\rangle_{\rho_{\nu_0}^{\mathrm{can}}\to\Gamma} + \langle W_{\mathrm{irr}}^\dagger\rangle_{\Gamma^*\to}^\dagger\right)$$
$$+ \frac{\beta}{2}\left(\langle W_{\mathrm{irr}}\rangle_{\rho_{\nu_0}^{\mathrm{can}}\to\Gamma^*} + \langle W_{\mathrm{irr}}^\dagger\rangle_{\Gamma\to}^\dagger\right) + O(\epsilon^3). \quad (35)$$

Here, by setting $f = \epsilon^{-1}W_{\mathrm{irr}}$ in (24), we have $\langle W_{\mathrm{irr}}\rangle_{\rho_{\nu_0}^{\mathrm{can}}\to\Gamma} + \langle W_{\mathrm{irr}}^\dagger\rangle_{\Gamma^*\to}^\dagger = O(\epsilon^2)$, where it should be noted that $W^* = -W^\dagger$. We thus obtain

$$\int d\Gamma \rho_\epsilon(\Gamma)(\langle W_{\mathrm{irr}}\rangle_{\rho_{\nu_0}^{\mathrm{can}}\to\Gamma} + \langle W_{\mathrm{irr}}^\dagger\rangle_{\Gamma^*\to}^\dagger)$$
$$= \int d\Gamma \rho_\epsilon(\Gamma^*)(\langle W_{\mathrm{irr}}\rangle_{\rho_{\nu_0}^{\mathrm{can}}\to\Gamma} + \langle W_{\mathrm{irr}}^\dagger\rangle_{\Gamma^*\to}^\dagger) + O(\epsilon^3)$$
$$= \int d\Gamma \rho_\epsilon(\Gamma)(\langle W_{\mathrm{irr}}\rangle_{\rho_{\nu_0}^{\mathrm{can}}\to\Gamma^*} + \langle W_{\mathrm{irr}}^\dagger\rangle_{\Gamma\to}^\dagger) + O(\epsilon^3) \quad (36)$$

This leads to the cancellation of the second and third lines in the right-hand side of (35) and yields the result

$$\mathcal{F}^{\mathrm{sym}}(\rho_\epsilon) - F_{\nu_1} = \frac{1}{2}\left(\langle W_{\mathrm{irr}}\rangle_{\rho_{\nu_0}^{\mathrm{can}}\to} - \langle W_{\mathrm{irr}}^\dagger\rangle_{\rho_\epsilon\to}^\dagger\right) + O(\epsilon^3). \quad (37)$$

From this relation, we also find that $\mathcal{F}$ in (17) and (29) and $\mathcal{S}$ in (30) are replaced by $\mathcal{F}^{\mathrm{sym}}$ and $\mathcal{S}^{\mathrm{sym}}$, respectively.

Concluding remarks: As a supplemental claim of this Rapid Communication, we explain the inequalities associated with the Sivak and Crooks relation (17). First, by recalling (10) and applying the non-negativity of relative entropy to (27) and (17), we obtain

$$\langle W_{\mathrm{irr}}^\dagger\rangle_{\rho_\epsilon\to}^\dagger \leq \langle W_{\mathrm{irr}}\rangle_{\rho_{\nu_0}^{\mathrm{can}}\to} + O(\epsilon^3), \quad (38)$$

and

$$\langle W^\dagger\rangle_{\rho_\epsilon\to}^\dagger \leq \langle W^\dagger\rangle_{\rho_{\nu_1}^{\mathrm{can}}\to}^\dagger + O(\epsilon^3), \quad (39)$$

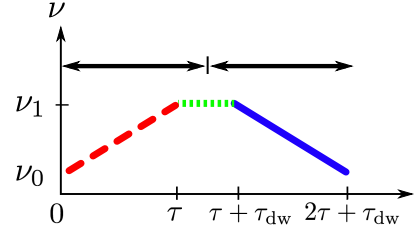respectively. These do not appear to be derived from the standard second law of thermodynamics. Furthermore,

FIG. 2: (Color online) Protocol in $t \in [0, \tau + \tau_{\mathrm{dw}}/2]$ consisting of $\hat{\nu}$ (dotted red line) and a dwell time in $[\tau, \tau + \tau_{\mathrm{dw}}/2]$, and the subsequent reverse protocol in $t \in [\tau + \tau_{\mathrm{dw}}/2, 2\tau + \tau_{\mathrm{dw}}]$, consisting of a dwell time in $[\tau + \tau_{\mathrm{dw}}/2, \tau + \tau_{\mathrm{dw}}]$ and $\hat{\nu}^\dagger$ (blue solid line).

we can generalize the inequality (39). Concretely, we consider protocols with dwell time τ_{dw} before starting the reversed protocol, as shown in Fig. 2. Let $\rho_\epsilon^{\tau_{\mathrm{dw}}}$ be the probability density at time $\tau + \tau_{\mathrm{dw}}$. Since $\rho_\epsilon^0 = \rho_\epsilon$ and $\rho_\epsilon^\infty = \rho_{\nu_1}^{\mathrm{can}}$, the work $\langle W^\dagger\rangle_{\rho_\epsilon^{\tau_{\mathrm{dw}}}\to}^\dagger$ is a generalization of the two works, $\langle W^\dagger\rangle_{\rho_\epsilon\to}^\dagger$ and $\langle W^\dagger\rangle_{\rho_{\nu_1}^{\mathrm{can}}\to}^\dagger$, which appear in (17). Now, we present the inequality

$$\langle W^\dagger\rangle_{\rho_\epsilon^{\tau_{\mathrm{dw}}}\to}^\dagger \leq \langle W^\dagger\rangle_{\rho_\epsilon^{\tau'_{\mathrm{dw}}}\to}^\dagger + O(\epsilon^3) \quad (40)$$

for any $\tau_{\mathrm{dw}} \leq \tau'_{\mathrm{dw}}$. The essence of the proof is to notice the monotonically decreasing property of the relative entropy in relaxation processes [30], that is,

$$\mathcal{D}(\rho_\epsilon^{\tau_{\mathrm{dw}}}|\rho_{\nu_1}^{\mathrm{can}}) \geq \mathcal{D}(\rho_\epsilon^{\tau'_{\mathrm{dw}}}|\rho_{\nu_1}^{\mathrm{can}}) \quad (41)$$

for any $\tau_{\mathrm{dw}} \leq \tau'_{\mathrm{dw}}$. In order to utilize this property, we define a new forward protocol in the time interval $[0, \tau + \tau_{\mathrm{dw}}/2]$ by combining $\hat{\nu}$ with no operation in a dwell time $\tau_{\mathrm{dw}}/2$, as shown in Fig. 2. For this protocol, we consider the Sivak and Crooks relation, and apply (41) to the expression of the relation. We then obtain (40). It should be noted that these inequalities are valid only for over-damped cases.

In sum, we have re-derived the Sivak and Crooks relation (17) by a simpler argument. We have emphasized that the relation is a non-stationary version of the extended Clausius relation proposed in the framework of SST. By considering a correspondence with SST, we successfully extended the Sivak and Crooks relation to (37) for under-damped cases. Although these relations are valid only near equilibrium, an exact relation can be easily derived from (21) without any approximation. The next important study will be to demonstrate how these relations are useful for understanding non-equilibrium phenomena.

[1] D. J. Evans, E. G. D. Cohen, and G. P. Morriss, Phys. Rev. Lett. **71**, 2401 (1993).
[2] G. Gallavotti and E. G. D. Cohen, Phys. Rev. Lett. **74**, 2694 (1995).
[3] J. Kurchan, J. Phys. A: Math. Gen. **31**, 3719 (1998).
[4] C. Maes, J. Stat. Phys. **95**, 367 (1999).
[5] G. M. Wang, E. M. Sevick, E. Mittag, D. J. Searles, and D. J. Evans, Phys. Rev. Lett. **89**, 050601 (2002)
[6] U. Seifert, Phys. Rev. Lett. **95**, 040602 (2005)
[7] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
[8] G. E. Crooks, Phys. Rev. E **61**, 2361 (2000).
[9] J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco, Jr., and C. Bustamante, Science **296**, 1832 (2002),
[10] D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco Jr., and C. Bustamante, Nature **437**, 231 (2005)
[11] R. Landauer, Phys. Rev. A **18**, 255 (1978).
[12] Y. Oono and M. Paniconi, Prog. Theor. Phys. Suppl. **130**, 29 (1998).
[13] T. Hatano and S. Sasa, Phys. Rev. Lett. **86**, 3463 (2001).
[14] E. H. Trepagnier, C. Jarzynski, F. Ritort, G. E. Crooks, C. J. Bustamante, and J. Liphardt, Proc. Natl. Acad. Sci. U.S.A. **101**, 15038 (2004).
[15] D. Ruelle, Proc. Natl. Acad. Sci. U.S.A. **100**, 3054 (2003).
[16] S. Sasa and H. Tasaki, J. Stat. Phys. **125**, 125 (2006).
[17] T. S. Komatsu, N. Nakagawa, S. Sasa, and H. Tasaki, Phys. Rev. Lett., **100**, 230602 (2008).
[18] T. S. Komatsu, N. Nakagawa, S. Sasa, and H. Tasaki, J. Stat. Phys. **142**, 127 (2011).
[19] G. Verley, R. Chétrite, and D. Lacoste, Phys. Rev. Lett. **108**, 120601 (2012)
[20] N. Nakagawa, Phys. Rev. E **85**, 051115 (2012).
[21] C. Maes and K. Netočný, arXiv:1206.3423.
[22] E. Bertini, D. Gabrielli, G. Jona-Lasinio, and C. Kadim, arXiv:1208.1872
[23] A. Mounier and A. Naert, arXiv:1208.4039.
[24] D. A. Sivak and G. E. Crooks, Phys. Rev. Lett. **108**, 150601 (2012).
[25] B. Gaveau and L.S. Schulman, Phys. Lett. A **229**, 347 (1997).
[26] K. Sekimoto, J. Phys. Soc. Jpn. **66**, 1234 (1997).
[27] K. Sekimoto, *Stochastic Energetics* (Springer, Berlin, 2010),
[28] C. Jarzynski, J. Stat. Phys. **98**, 77 (2000).
[29] R. Kawai, J. M. R. Parrondo, and C. Van den Broeck, Phys. Rev. Lett. **98**, 080602 (2007).
[30] T. M. Cover and J. A. Thomas: *Elements of Information Theory, 2nd edition* (John Wiley and Sons, Now York, 2006).
[31] T. S. Komatsu and N. Nakagawa, Phys. Rev. Lett. **100**, 030601 (2008).
[32] T. S. Komatsu, N. Nakagawa, S. Sasa, and H. Tasaki, J. Stat. Phys. **134**, 401 (2009).